SFRP2/DPP4 and FMO1/LSP1 Define Major Fibroblast Populations in Human Skin

Q18 Tracy Tabib[1], Christina Morse[1], Ting Wang[2], Wei Chen[2] and Robert Lafyatis[1]

Fibroblasts produce matrix, regulate inflammation, mediate reparative processes, and serve as pluripotent mesenchymal cells. Analyzing digested normal human skin by single-cell RNA sequencing, we explored different fibroblast populations. T-distributed stochastic neighbor embedding and clustering of single-cell RNA sequencing data from six biopsy samples showed two major fibroblast populations, defined by distinct genes, including *SFRP2* and *FMO1*, expressed exclusively by these two major fibroblast populations. Further subpopulations were defined within each of the *SFRP2* and *FMO1* populations and five minor fibroblast populations, each expressing discrete genes: *CRABP1, COL11A1, FMO2, PRG4,* or *C2ORF40*. Immunofluorescent staining confirmed that *SFRP2* and *FMO1* define cell types of dramatically different morphology. *SFRP2*$^+$ fibroblasts were small, elongated, and distributed between collagen bundles. *FMO1*$^+$ fibroblasts were larger and distributed in both interstitial and perivascular locations. Differential gene expression by *SFRP2*$^+$, *FMO1*$^+$, and *COL11A1*$^+$ fibroblasts suggests roles in matrix deposition, inflammatory cell retention, and connective tissue cell differentiation, respectively.

Journal of Investigative Dermatology (2017) ■, ■–■; doi:10.1016/j.jid.2017.09.045

INTRODUCTION

Fibroblasts secrete extracellular matrix, mediating reparative and fibrotic processes. Fibroblasts play key roles in healing wounds but also as the mediators of fibrosis. In addition, they regulate inflammation by anchoring leukocytes and regulating immune cell functions. In lymphoid tissues they contribute to secondary lymphoid organ structure, and in nonlymphoid tissues to the development of tertiary lymphoid structures (Barone et al., 2016). In carcinogenesis fibroblasts have an emerging role, because cancer-associated fibroblasts support and regulate tumor cell growth (Ohlund et al., 2014). Thus, fully understanding the complexity of normal dermal fibroblasts is key to understanding their roles in a wide variety of pathological conditions.

Dissecting fibroblast functional heterogeneity has lagged behind understanding of inflammatory cell types because of a lack of discrete markers and limitations of technologies such as fluorescence-activated cell sorting on enzymatically digested tissues. Thus, we have no comprehensive understanding of the repertoire of tissue fibroblasts. Because of skin accessibility and complexity as a connective tissue, dermal mesenchymal cells have been a focus of several groups. Watt's group has described different well-defined cell populations in normal skin. These studies have shown that arrector pili/smooth muscle cells selectively express ITGA8 and NPNT (Fujiwara et al., 2011). Furthermore, this group has defined a series of markers that are stable or dynamic for dermal papilla (CRABP1), papillary (DPP4/CD26), and reticular (PDPN, SCA1/ATXN1) fibroblasts (Driskell et al., 2013; Driskell and Watt, 2015).

Recent murine studies have advanced several alternative approaches to understanding mesenchymal cell heterogeneity in various tissues and of profibrotic cell progenitors. One approach to understanding mesenchymal cell heterogeneity is single-cell cloning. A recent study showed that different functional, clonal populations could be characterized on the basis of HAS2/MMP10 versus COL1A1/DCN/MMP2 expression (Hiraoka et al., 2016). However, a major limitation to cell cloning is uncertainty regarding the stability of cell phenotypes on expansion of cells in vitro. In addition, it may be difficult to identify rare, discrete mesenchymal cell types.

Several groups have begun to examine cellular heterogeneity in mouse embryonic mesenchymal cells directly using single-cell analyses. Singhal et al. (2016), analyzing cultured embryonic mesenchymal cells by flow cytometry, distinguished six subpopulations of cells based on CD73 (NTSE), CD146 (MCAM), CD90, PDPN, CD24, and CD38. This heterogeneity was stable for 4û6 passages. Rinkevich et al. (2015), using single-cell microfluidic gene expression analysis, showed that a population of embryonic lineage fibroblasts purified from adult murine skin expresses COL1A1, COL3A1, FBN1, PDGFRA, VIM, DCN, S100A4, and other genes (Rinkevich et al., 2015). Engrailed-positive versus engrailed-negative cell types could be further distinguished based on expression of other sets of genes (transcriptomes). Most significantly, mice in which engrailed-derived cells were deleted showed less scarring after wound healing. In addition, DPP4 was shown to be a marker of engrailed-positive cells, and its inhibition led to decreased wound scarring. These murine studies have clarified the importance

[1]*Division of Rheumatology and Clinical Rheumatology, Department of Medicine, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania, USA; and* [2]*Division of Pulmonary Medicine, Allergy and Immunology, Department of Pediatrics, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania, USA*

Correspondence: Robert Lafyatis, 200 Lothrop Street, BST S720, Pittsburgh, Pennsylvania 15260, USA. E-mail: lafyatis@pitt.edu

Abbreviations: IF, immunofluorescence; RNA-seq, RNA sequencing; SLM, smart local moving; t-SNE, t-distributed stochastic neighbor embedding

Received 21 July 2017; revised 31 August 2017; accepted 23 September 2017; accepted manuscript published online XXX; corrected proof published online XXX

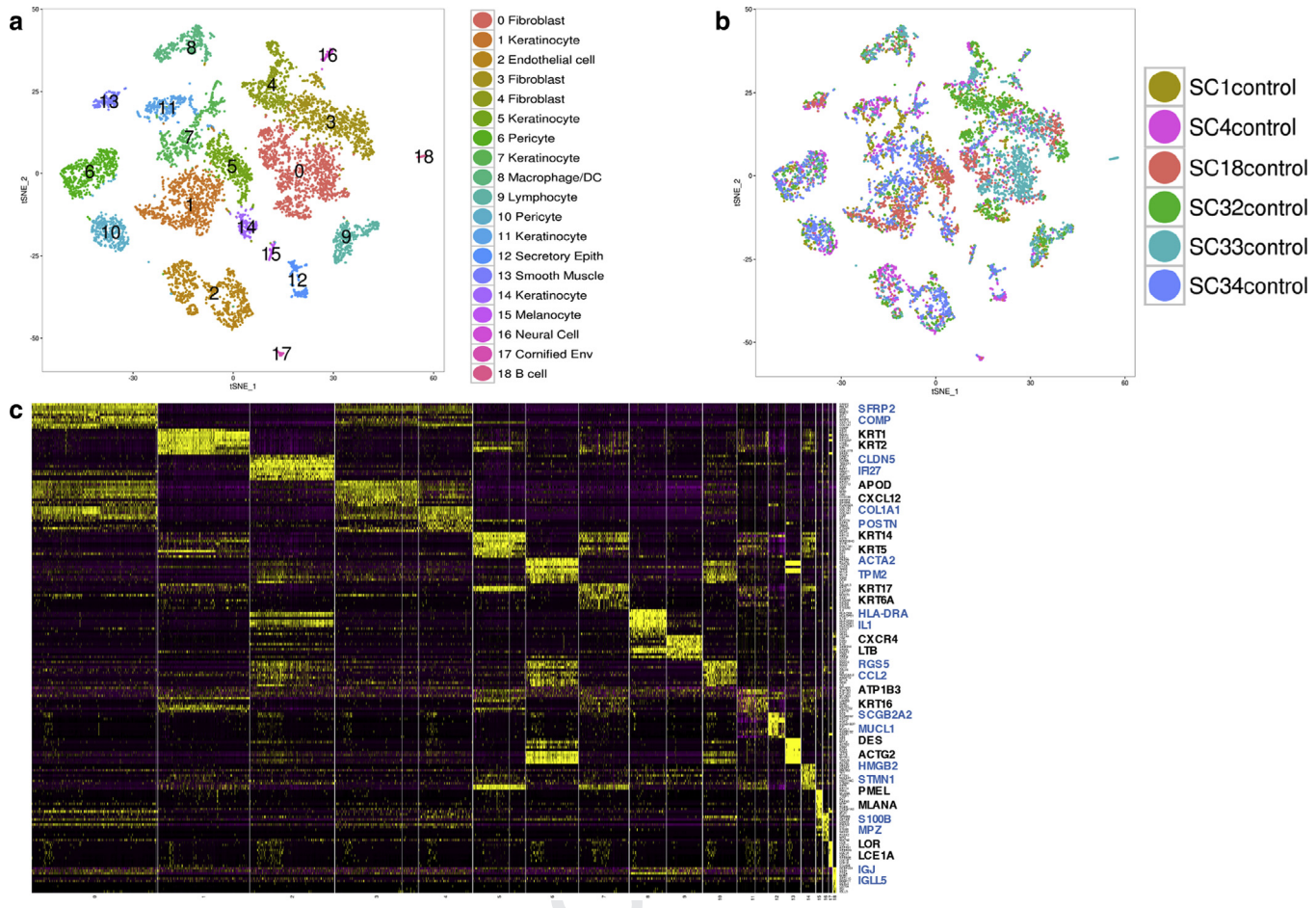

Figure 1. t-SNE shows groupings and SLM clustering of normal skin cell populations. (**a**) Each point represents a cell. Two-dimensional t-SNE shows dimensional reduction of cell transcriptomes. Cells are colored by K-nearest neighbors graph based on Euclidean distance in principal component analysis space using an SLM algorithm to iteratively group cells. (**b**) Cells are grouped by t-SNE as in **a** but are colored according to the subject identity. (**c**) Transcriptomes of 8,522 cells from six normal skin biopsy samples clustered using Seurat (SLM clustering). Each column represents a cell. The five genes most differentially expressed between each cluster are shown, and two of these five genes are enlarged to help identify each cluster. Cluster numbers, indicated at the bottom, are as shown in **a**, t-SNE. SLM, smart local moving; t-SNE, t-distributed stochastic neighbor embedding.

of understanding fibroblast heterogeneity in skin and have provided several transcriptome datasets that can aid in identifying fibroblast subpopulations in human skin.

Using advances in single-cell RNA sequencing (RNA-seq) technology, allowing thousands of single cells to be analyzed in a single experiment, we examined single-cell transcriptomes of cell populations from whole skin without prepurifying fibroblast populations. We identify multiple discrete dermal fibroblast populations, including two major and five minor fibroblast types, strongly suggesting underlying functional heterogeneity.

RESULTS

Detection of skin transcriptomes corresponding to epithelial, endothelial, and mesenchymal cell types in normal skin

Using single-cell RNA-seq, we examined gene expression in all cells obtained from enzymatically digested skin from six healthy control skin samples. Dorsal mid-forearm skin samples were analyzed from both male and female subjects of varying ages (see Supplementary Table S1 online). We examined a total of 8,522 cells from six subjects (1,135–1,748 cells/sample) (see Supplementary Table S1). To gain power to detect rare cell types and to examine the reproducibility of cell types between subjects, cell transcriptomes from all the samples were grouped and analyzed together.

Combined cell-gene count matrices were analyzed by principal component analysis, and the statistical significance of principal components was analyzed using Jackstraw (Chung and Storey, 2015). Statistically significant principal components were used for t-distributed stochastic neighbor embedding (t-SNE) dimensional reduction and visualization and for clustering. Cells were clustered using an unsupervised graph-based clustering algorithm (smart local moving [SLM] clustering, described in the Methods section), which in total identified 19 distinct clusters of cells, distinguished by color (Figure 1a). Cells from each subject were also indicated by different colors (Figure 1b), showing that each cluster included cells from each biopsy sample. SLM clusters contained genes well known to be expressed by various cell types in the skin, permitting the cells in each cluster to be identified (Figures 1c and 2). DES clearly identified a cell cluster of smooth muscle cells; KRT1 and KRT14 as clusters